3

The WSRT CVn H I mass function

We construct the H I mass function (H IMF) from the results of the Westerbork Synthesis Radio Telescope (WSRT) Canes Venatici (CVn) survey. The WSRT CVn survey is a new blind extragalactic survey in the 21-cm line. The survey covers an area of approximately 86 deg^2 and an approximate redshift interval $-450 < cz < 1330$ km s^{-1}. We detected 70 H I sources in the surveyed area, of which one detection does not have an obvious optical counterpart. The constructed H IMF has a slope of -1.17, which falls in the range of the slopes of the H IMF derived from previous H I surveys. We address the influence of the uncertainties in the adopted distances and the measured integrated fluxes on the estimated slope using Monte Carlo simulations. Adding realistic uncertainties to distances and to the integrated fluxes, one can change the slope of the WSRT CVn H IMF by 0.02 and -0.04, respectively. Assuming the H I Parkes All Sky Survey (HIPASS, Barnes et al. 2001) value for the kink in the H IMF, the slope will steepen by -0.05. The H IMF obtained from the results of the WSRT CVn survey is accurate down to $\log(h_{70}^2 M_{\rm HI}/M_\odot) = 6.4$, which is the lowest H I mass reached by any of the previous H IMF estimates.

3.1 Introduction

One of the goals of galaxy surveys is to describe the distribution and statistical properties of galaxies in the Universe, and ideally, to describe the underlying mass distribution. A translation from the observed emission to the mass of the material can be extremely complex, as it is a function of various physical and chemical parameters. Under most conditions, the physics responsible for the 21-cm emission line of H I does not depend on the environmental parameters (see Kulkarni & Heiles 1988; Schneider 1996). When carrying out an observation in the 21-cm line, it is straightforward to assign a mass which corresponds to the measured flux.

The luminosity function is a statistical tool to describe the distribution of luminosities of galaxies measured in a unit volume in a bin with a certain luminosity (or magnitude). The H I mass function (H IMF) is defined analogously to the luminosity function - it gives the number of objects with measured H I masses per unit volume in a bin of a certain H I mass. Conceptually, the H IMF was introduced by Briggs (1990) in order to test the completeness of H I observations of sources selected from optical catalogues. The H IMF was derived from the optical luminosity function, using an assumed relation between the H I content and the luminosity of galaxies. Since then, the H IMF has been derived using a variety of H I data: H I observations of optically selected galaxies (e.g. Briggs & Rao 1993; Solanes et al. 1996; Springob et al. 2005), as well as using H I detections revealed from blind H I surveys (e.g. Zwaan et al. 1997; Rosenberg & Schneider 2002; Zwaan et al. 2005b). While the largest blind H I surveys cover a range of environments, most of the surveys were targeted to have a census of H I detections in a specific environment. A blind H I survey carried out by Verheijen et al. (2000, 2001) has been conducted in the Ursa Major (UMa) cluster, while the H I survey of Szomoru et al. (1996) has been targeted on infrared selected galaxies in the Boötes void.

Since the introduction of the H IMF as a tool, the major conclusions drawn from the H IMF have not changed much over the past years. No large population of low surface brightness (LSB) galaxies that possibly has been missed in optical surveys appears to exist, neither (or at least very few) a self-gravitating H I cloud has been detected which did not undergo any star formation process (Zwaan et al. 2005b). However, the number of objects with small H I masses, predicted by the slope of the H IMF, has been poorly constrained. The slope α of the H IMF can be as shallow as $\alpha \approx$ -1 (Verheijen et al. 2000, 2001) or as steep as $\alpha \approx$ -1.5 (Henning et al. 2000).

A number of reasons have been discussed in the literature as a cause for this discrepancy: low-number statistics for objects with small H I masses, distance uncertainties, effects of the environment, as well as differences in the methodology and in addressing of completeness issues of the surveys. Due to the weakness of H I photons in combination with the limitations of current radio telescopes, surveys carried out in the 21-cm line cover generally much smaller and shallower volumes than surveys conducted in optical or infrared wavelengths. While recent redshift surveys, such as the Two Degree Field Galaxy Redshift Survey (2dFGRS, Colless et al. 2001) and the Sloan Digital Sky Survey (SDSS, York et al. 2000), provide measurements of hundreds of thousands of galaxies for statistical studies, the largest H I survey up to date, H I Parkes All Sky Survey (HIPASS, Barnes et al. 2001) contains 4315 objects detected in the whole southern sky with $\delta \leq$ 2 deg. It may be difficult to pin-point down one reason, or a combination of several reasons, for the measured spectrum of the H IMF slopes, because the overall number

of H I sources detected in the (HI) surveys is relatively small. This makes a variety of additional statistical tests not applicable. On the other hand, the sizes of optical and infrared surveys make it possible to study the additional dependence of the luminosity function on various physical parameters. It has been undoubtfully shown that the slope of the optical luminosity function becomes steeper as one goes from passive to active star forming galaxies (Madgwick et al. 2002). Similarly, the slope becomes steeper as one moves from redder to the bluer galaxies (Blanton et al. 2001). Up to now, only two studies of the H IMF are derived from samples of H I detections that are large enough to be divided into a few subsamples with given criteria. The slope of the H IMF steepens when moving from earlier towards later type galaxies (Zwaan et al. 2003; Springob et al. 2005). A trend in the slope of the H IMF with the environment is more uncertain (Zwaan et al. 2003; Springob et al. 2005).

The discrepancy at the faint end of the H IMF was one of the main scientific drivers for conducting the Westerbork Synthesis Radio Telescope (WSRT) Canes Venatici (CVn) survey, a blind extragalactic 21-cm survey sensitive enough to detect objects of H I mass of 10^6 $M_\odot$ at a distance of 4.6 Mpc (based on Monte Carlo simulations described in Section 2.4 in Chapter 2). This survey will tie down the faint end of the H IMF with a high significance between H I masses of 5×10^6 and 10^8 $M_\odot$.

The H IMF of the WSRT CVn survey is the topic of this chapter. In Section 3.1 we briefly describe the WSRT CVn survey and some of the properties of the detected objects. In Section 3.3 we discuss first the usage of different models for obtaining the distances to objects from their measured redshifts. Second, we calculate the H I masses using the adopted distances and present a few relations between the H I mass and other parameters. In Section 3.4 we estimate the H IMF and discuss the uncertainties in the slope of the H IMF. We compare the WSRT CVn H IMF with various literature H IMFs in Section 3.5. Discussion on the discrepancy of the H IMF slopes from the WSRT CVn survey and previous surveys is presented in Section 3.6. The final conclusions are presented in Section 3.7.

3.2 Summary of the WSRT CVn blind H I survey

We have carried out a blind H I survey in the CVn region covering an area of about 86 deg^2 and a velocity range of approximately -450 to 1330 km s^{-1}. The observations of approximately 60 $\times$ 12 hr have been conducted with the WSRT during 2001, 2002 and 2004. The survey and the data reduction process, the search procedure and the determination of the properties of the detected objects as well as properties by themselves are presented in Chapter 2. We summarise the survey parameters in Table 3.1. In the following paragraph we recapitulate the most important steps in the data analysis and we give an overview of the detections obtained.

The observed data have been reduced using scripts based on the MIRIAD programmes (Sault et al. 1995) and programmes written for this project (T.A.O. and K.K). From each of the observed pointings we produced a three dimensional datacube. We created the final line datacubes by extracting the central part of the line datacubes made by combining the individual datacubes with centres less than 22 arcmin apart. Datacubes were produced in five different resolutions: the datacubes reduced were smoothed twice in the spatial (Ra, Dec) domain and twice in the velocity (cz) domain independently,

Parameter	WSRT CVn value
Dec range	$\delta_1 = 31^0 33^{'} 00^{''}$ $\delta_2 = 46^0 18^{'} 00^{''}$
Ra range for δ_1	$12^h 19^m 55.2^s \leq \alpha \leq 12^h 47^m 2.4^s$
Ra range for δ_2	$12^h 16^m 42.9^s \leq \alpha \leq 12^h 50^m 3.0^s$
Velocity range	-450 $\leq cz \leq$ 1330 km s^{-1}
Velocity resolution	33 km s^{-1}
Channel separation	16.7 km s^{-1}
Spatial resolution (FWHM beamwidth)	30 $\times$ 60 arcsec
Integration time per pointing	$\approx$ 80.1 min
H I mass limit (from simulations)	4.7×10^4 $distance^2$ $M_\odot$
Number of detections	70

Table 3.1: WSRT CVn survey summary.

to allow detection of objects with lower signal-to-noise ratios. We used the number distributions of the sizes of regions, made from connected pixels above a certain signal to noise ratio, to define the maximum sizes of noise structures in the line datacubes. The search process was based on the detection of a region of connected pixels with a minimum number of connected pixels larger by at least one pixel than the defined maximum size of a noise structure for that particular resolution and signal-to-noise ratio of the datacubes searched. This search process applied to the whole survey (1372 line datacubes at 5 different resolution) revealed 70 H I detections. From the 70 H I detected objects, only one (WSRT-CVn-61) does not have an optical counterpart. This object is close to NGC4288 and has been detected in H I before (Wilcots et al. 1996). The other 69 detections can be cross-correlated with optically identified galaxies using their positions on the sky and, if available, their velocities. Four of those galaxies do not have any reliable entry in existing optical or infrared catalogues. We do, however, find their optical counterparts on the (X)DSS images. From a comparison with published H I observations we conclude that 19 of the objects have been detected for the first time in the 21-cm emission line.

We have carried out a Monte Carlo simulation to describe the uncertainties of the derived parameters and the completeness of the survey. From a comparison with previously published H I observations and the WSRT follow-up of H I detections selected from the blind survey, we conclude that the reliability (percentage of confirmed sources) of the WSRT CVn survey is 100%. MAPS_NGP O_218_0783987 is the only object in the survey volume which has been detected in H I previously (Huchtmeier et al. 2000), but not in the WSRT CVn survey. This object falls in the range of properties for which the WSRT CVn survey is slightly incomplete (C > 90%). In the WSRT CVn data there is no sign of any H I emission at the position of MAPS_NGP O_218_0783987. Based on our synthesis imaging data of high quality, we are inclined to believe that the previous H I detection of this object is not real.

3.3 H I masses of objects detected in the blind WSRT CVn survey

The H I mass of an object can be calculated using the relation

$$M_{HI}/M_{\odot} = 2.356 \times 10^5 D^2 S_{int}, \tag{3.1}$$

where D is the distance to the object in Mpc and S_{int} is the integrated flux in Jy km s^{-1}. This relation is valid under the assumption that the H I is optically thin.

3.3.1 Distances to the objects

Until now, only a small number of galaxies is known (~ 600) for which distances have been measured with high precision, using primary distance indicators. For about an additional 5000 objects, distances have been estimated using secondary distance estimators (e.g. Masters et al. 2004). For all other objects, redshift provides the only estimate of their distance.

With H I synthesis observations, both the spatial and the redshift information of an object are collected at the same time. The major problem remains how to estimate an accurate distance from the measured redshift. The measured recession velocities deviate from the pure Hubble flow with a typical value for the peculiar velocity reaching up to a few hundred kilometres per second. So, for low recession velocities the derived distances can be very uncertain. The nearby Universe (up to 10 Mpc) is the only place where current radio telescopes are sensitive enough to detect small amounts of H I and here, peculiar velocities play an important role. The H I masses of the nearest objects can be seriously and systematically over– or underestimated, if one neglects the influence of peculiar velocities on the measured recession velocity. This uncertainty may cause a change in the slope of the H IMF.

It is predicted that the deviation of observed recession velocities from the Hubble flow (peculiar velocity) in the Local Universe is caused by the gravitational deceleration by the mass of the Local Group (Lynden-Bell 1981; Sandage 1986), while at larger distances the deviation can be caused by the gravitational influence of nearby groups and the Virgo-centric flow. From observations it has been known for more than three decades that the Hubble flow just outside the Local Group is dynamically cold (Sandage et al. 1972). The conclusion on the coldness of the local flow follows from the small scatter in the velocity-distance relation of nearby galaxies. For galaxies which lie in the range of distances 5-7 h^{-1} Mpc, the line-of-sight peculiar velocity dispersion σ_H is in the range 80-40 km s^{-1} (Giraud 1986; Schlegel et al. 1994; Sandage 1999; Ekholm et al. 2001; Karachentsev et al. 2003a). The velocity dispersion increases with distance and Macciò et al. (2005) provide the best fit to this dependence

$$\sigma_H = 88 \pm 20 \text{ kms}^{-1} \times (D/7\text{Mpc}), \tag{3.2}$$

for distances less than 10 Mpc. Averaged over larger volumes, the velocity dispersion increases to a few hundred km s^{-1}. More precisely, in a volume of radius of 14 h^{-1} Mpc the velocity dispersion has a value of 200 km s^{-1} (Groth et al. 1989), while in the volume of radius of 30 h^{-1} Mpc it increases up to 310 km s^{-1} (Groth et al. 1989; Tonry et al. 2000).

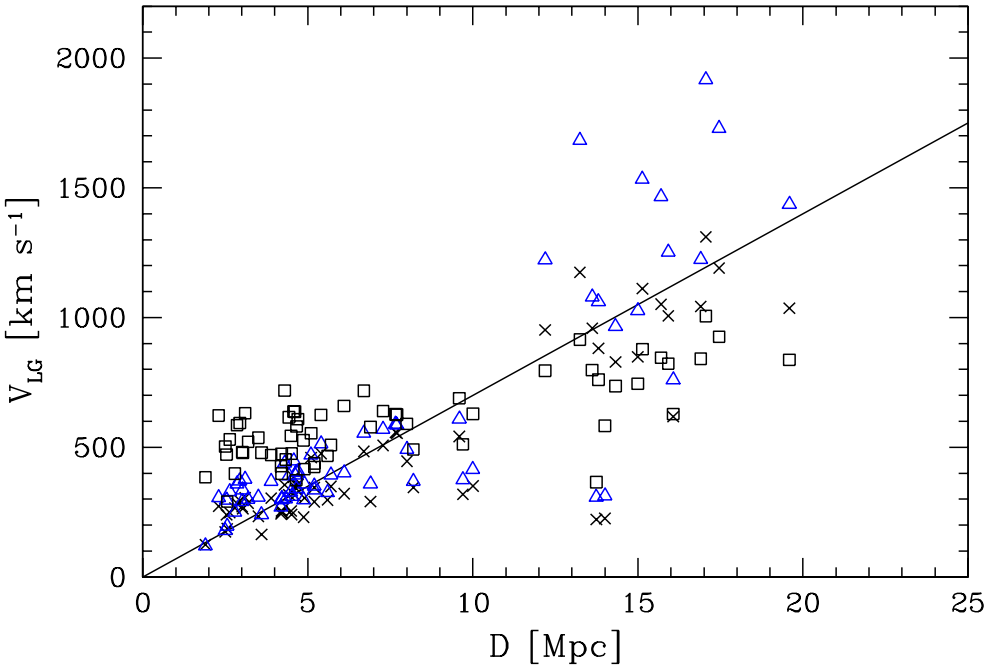

Figure 3.1: Velocity-distance diagram for 64 galaxies in the CVn constellation area. Distances D are estimated independent of the observed recession velocities. Crosses correspond to recession velocities measured in the Local Group frame V_{LG}, triangles correspond to V_{LG} velocities corrected for the peculiar velocities interpolated from the grid provided by S. Zaroubi (private communication), squares correspond to V_{LG} velocities corrected for peculiar velocities using the multi–attractor velocity model by Tonry et al. (2000). The line corresponds to a Hubble flow with $H_0 = 70$ km s^{-1} Mpc^{-1}.

Under the assumption that galaxies are unbiased tracers of the large–scale velocity field caused by gravity, peculiar velocities are a direct measurement of the underlying mass distribution. Peculiar velocities can be calculated from the observed velocities, and they can be related to the density fluctuation field via gravitational–instability theory. Reconstruction of the density field and peculiar velocities can be carried out in a statistical sense (Branchini et al. 1999; Zaroubi et al. 1995; Romano-Diaz 2004), providing the resulting velocity and density field on a discretely spaced grid.

Another approach to measure the peculiar velocities is to construct a parametric model of the local velocity field, which clearly is an oversimplified way to calculate the peculiar velocity at a given point. The parametric models provide a direction-dependent redshift distance template, produced by adding the influence of the individual components (such as Hubble flow and mass-attractors) to the velocity flow independently (recent models are provided by Tonry et al. 2000, and Mould et al. 2000).

The nearby CVnI/II groups of galaxies, selected for the WSRT CVn survey, have been very well studied in the literature. The CVn groups extend along a line of sight with

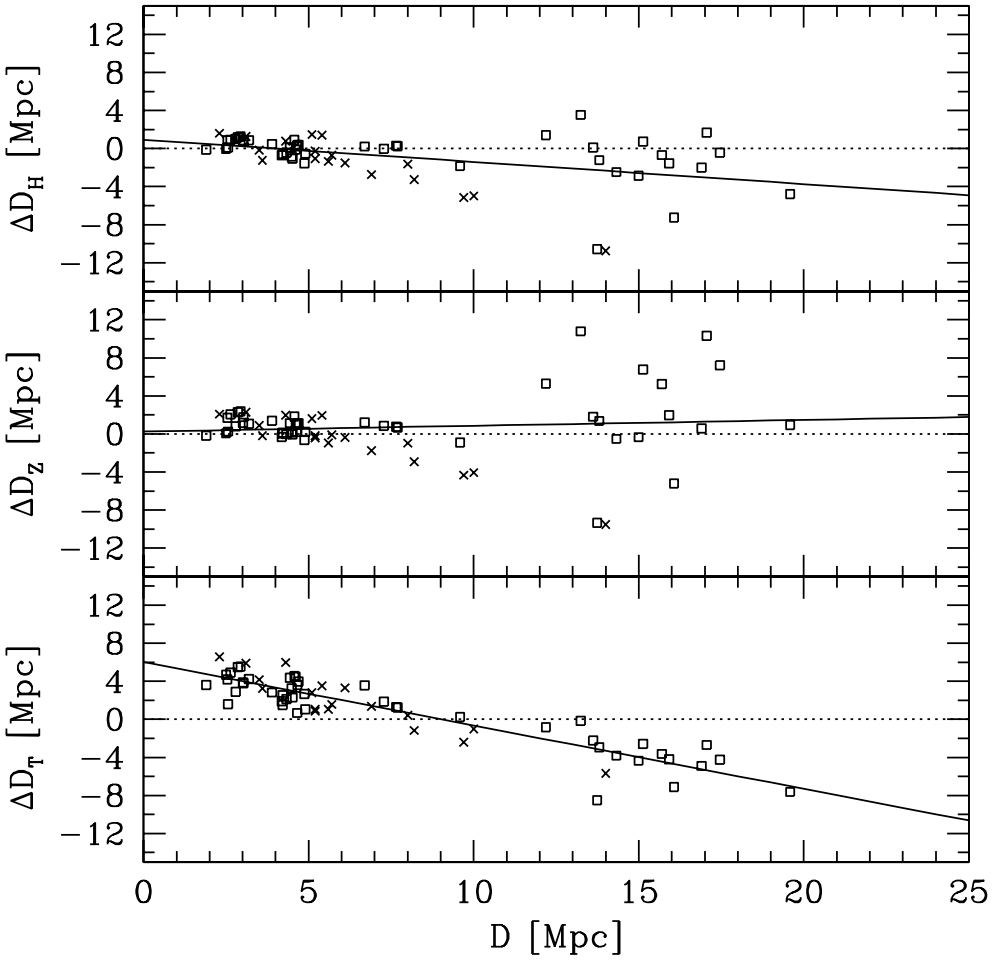

Figure 3.2: Differences between the distances estimated from three velocity models and the independently estimated distances D for galaxies in the CVn area. In the top panel, modelled distances have been calculated assuming pure Hubble flow, $D_H = V/H_0$, with $H_0 = 70$ km s^{-1} Mpc^{-1}. In the two lower panels distances have been calculated from the Hubble flow using recession velocities previously corrected by peculiar velocity. The peculiar velocities have been obtained from the velocity flow model of S. Zaroubi (middle panel) and from the Tonry et al. (2000) multi–attractor model. Galaxies for which distances have been estimated using one of the distance indicators with the accuracy of 5-15% are presented with the open squares. Crosses correspond to galaxies with the distances estimated with the accuracy of 20-30%. The continuous line is the least-square fit to all data points in each of the panels.

recession velocities spanning an interval from approximately 200 to 1200 km s^{-1}, and they are known to be rich in dwarf galaxies. Projected on the sky, the CVnI/II groups lie in the area of the CVn constellation ($\alpha = 11^h30^m$ to 13^h40^m, $\delta = +25^0$ to $+55^0$).

Karachentsev et al. (2003b) studied the velocity flow in the CVnI cloud. They collected 50 galaxies in the CVn area with velocities below 400 km s^{-1} for which distances have been measured independently and for which redshift measurements are available. We extend their sample with an additional 14 galaxies from the Surface Brightness Fluctuation (SBF) Survey (Tonry et al. 2001), which fall inside the area of the CVn constellation limits. In total there are 64 galaxies with independent distances: 17 galaxies with distances measured from SBF (Tonry et al. 2001), 29 galaxies with distances estimated using the tip of the red giant branch (TRGB) method (Karachentsev et al. 2002, 2003b), 17 galaxies with distances estimated from measurements of the brightest stars (BS, Makarova et al. 1998; Sharina et al. 1999) and one galaxy with a distance obtained from Cepheids (Sandage & Tammann 1982). Distances from the measurements based on the luminosity of the TRGB, Cepheids and SBF are the most reliable methods, giving distances with an accuracy of $\sim$ 5-15%. A characteristic error for distances estimated using the BS method is $\sim$ 20-30% (Karachentsev et al. 2003a).

We do not intend here to model the velocity field in the area of the CVn groups of galaxies. Instead, our plan is to compare the different ways of calculating distances to the detected galaxies and select the method with the smallest uncertainties. The selected method will be used to calculate distances to the objects detected in the WSRT CVn survey.

We consider two methods to estimate the peculiar velocity of an object and correct the recession velocity measured for it (two distance models). First, we use the velocity field reconstructed from the peculiar velocities of the detections in the SBF survey (kindly provided by S. Zaroubi). The reconstruction method has been described in Zaroubi et al. (1995) and Zaroubi (2000). The reconstructed velocities (distances) and densities are given on a grid with a spacing of 0.75 h^{-1} Mpc. We use an iterative method to calculate the peculiar velocity (or the distance) of an object by interpolating the peculiar velocities given in the closest 8 grid points. We correct the recession velocity for the calculated amount of peculiar velocity and calculate a new distance to the object, till the differences in the peculiar velocities in two consecutive iterations are smaller than 1 km s^{-1}. Second, we use Tonry's et al. (2000) parametric model to obtain the peculiar velocities. This model is based on a maximum likelihood fit of the modelled relation connecting velocity and distance to the SBF survey data. The modelled relation includes a thermal velocity dispersion, a number of spherical attractors and an additional quadrupole correction to the Hubble flow. In addition, we calculate the distances from the measured recession velocities, assuming an unperturbed Hubble flow (the third distance model).

Recession velocities of 64 objects in the CVn area corrected for the peculiar velocity are presented in Figure 3.1 as a function of distance, in the Local Group (LG) frame. Crosses correspond to the recession velocities observed, triangles correspond to interpolated recession velocities obtained from the reconstructed velocity field (provided by S. Zaroubi) and squares correspond to velocities corrected for peculiar velocities obtained from the multi–attractor velocity model by Tonry et al. (2000), modified to use the assumed value $H_0 = 70$ km s^{-1} Mpc^{-1}. All velocities are given in the LG frame. It is clear that velocities corrected for non-Hubble motions using Tonry's et al. (2000) model show a systematic offset with respect to the line representing the relation $V = H_0 D$, with $H_0 = 70$ km s^{-1} Mpc^{-1}.

This offset is visible more clearly in Figure 3.2, where we show the difference between distances estimated from the recession velocity observed, corrected for the amount of

Method	σ_{ALL} [Mpc]	$\sigma_{Q=1}$ [Mpc]	$\sigma_{\mathrm{D}\leq 10\mathrm{Mpc}}$ [Mpc]	$\sigma_{\mathrm{D}>10\mathrm{Mpc}}$ [Mpc]
V_H	2.55	2.28	1.47(8)	4.14
V_Z	3.26	3.20	1.47(5)	6.03
V_T	3.61	3.78	1.95(0)	2.30

Table 3.2: Standard deviation of the differences in distances. The first row is based on calculations assuming that the V_{LG} velocities measured correspond to the Hubble flow. V_Z velocities, used for the calculations presented in the second row, have been obtained from recession velocities after correcting them for peculiar velocities from the velocity field given on a grid provided by S. Zaroubi. Results in the third row have been obtained using the recession velocities corrected for the peculiar velocity from the Tonry et al. (2000) multi–attractor model.

peculiar velocity estimated by one of the models, and distances estimated independently. It has to be kept in mind that the models and the data are far from being independent: both the grid provided by S. Zaroubi and the Tonry et al. (2000) model are based on the SBF survey data, and basically all of the detections in the CVn constellation area at distances larger than 15 Mpc are collected from the SBF survey.

We calculate the dispersion of differences between the distances derived from the three different models considered and the independently estimated distances. The calculated dispersion for all 64 galaxies is given in the first column in Table 3.2. In addition, we calculate the dispersion for the subsample of galaxies with distances obtained using one of the distance indicators with an accuracy of $\sim$ 5-15%. These values have been presented in the second column in Table 3.2 (marked with "Q=1"). Considering the generally larger uncertainties in the distances for nearby objects, we divide the sample in two subsamples with respect to a distance of 10 Mpc. The dispersion of differences between the modelled and real distances for galaxies at distances smaller or equal than 10 Mpc (with the average distance of 4.9 Mpc) and at distances larger than 10 (with the average distance of 15.2 Mpc) Mpc is given in the third and forth column in Table 3.2, respectively.

The multi–attractor model (Tonry et al. 2000) shows a severe systematic disagreement with respect to the Hubble flow with $H_0 = 70$ km s^{-1} Mpc^{-1}. We exclude this model from any future consideration because its systematic effects. The distances obtained from the model of S. Zaroubi follow the assumed Hubble flow very well. That is also the case for distances obtained from recession velocities with zero peculiar velocities - but only at distances smaller than about 10 Mpc.

In addition, we compared the dispersions in the differences between the actual distances and distances derived with various models of velocity flow. This dispersion is the smallest when using the model of an unperturbed Hubble flow. Given the small number of galaxies in the sample - there are only 16 galaxies at distances larger than 10 Mpc, we base our choice for the velocity model on these dispersions of differently estimated distances. We conclude, based on the three models explored, that the best way to obtain distances for the detections in the CVn region is to use the measured recession velocities in the LG frame, neglecting possible peculiar velocities.

WSRT CVn id	NED name	α h m s	δ d m s	cz km s^{-1}	Distance Mpc	Distance method	Reference of the method
2	UGC07698	12 32 54.4	+31 32 28	309	6.1	BS	Makarova et al. (1998)
8	UGCA292	12 38 40.1	+32 46 01	294	3.1	BS	Makarova et al. (1998)
14	NGC4395	12 25 48.9	+33 32 48	300	4.6	RGB	Karachentsev et al. (2003b)
21	UGC07605	12 28 38.9	+35 43 03	310	4.4	RGB	Karachentsev et al. (2003b)
23	UGC07949	12 46 59.8	+36 28 35	339	10.4	BS	Makarova et al. (1998)
24	UGC07559	12 27 05.1	+37 08 33	223	4.9	RGB	Karachentsev et al. (2003b)
26	UGC7599	12 28 28.6	+37 14 01	282	6.9	BS	Makarova et al. (1998)
33	IC3687	12 42 15.1	+38 30 12	367	4.6	RGB	Karachentsev et al. (2003b)
34	UGCA290	12 37 21.8	+38 44 38	471	6.7	RGB	Crone et al. (2002)
49	UGC07577	12 27 40.9	+43 29 44	236	2.5	RGB	Karachentsev et al. (2002)
53	NGC4460	12 28 45.6	+44 51 51	546	9.6	SBF	Tonry et al. (2001)
68	NGC4449	12 28 11.9	+44 05 40	244	4.2	RGB	Karachentsev et al. (2003b)
69	NGC4244	12 17 29.6	+37 48 26	255	4.5	RGB	Karachentsev et al. (2003b)

Table 3.3: Objects with independently estimated distances.

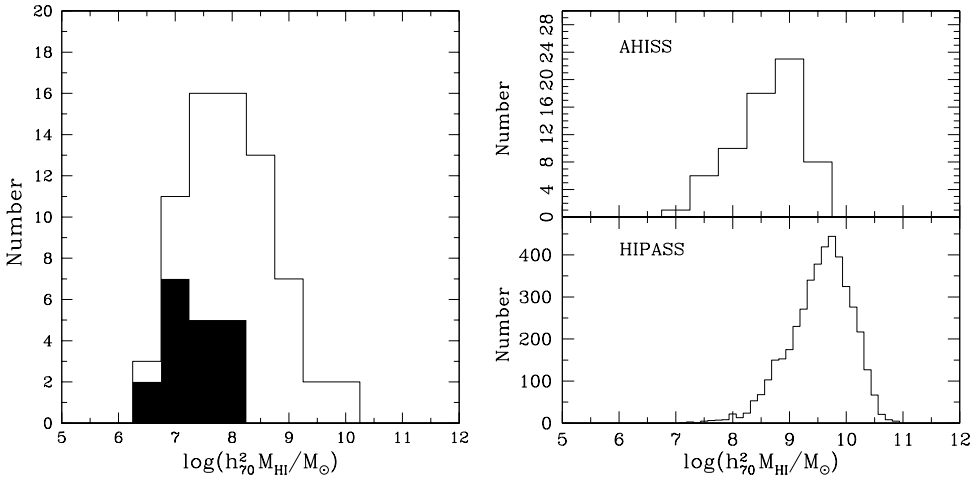

Figure 3.3: Number of detections per H I mass bin. Left: The empty histogram represents the distribution of the whole sample of 70 detections in the WSRT CVn survey, the filled histogram represents the distribution of 19 detections observed for the first time in H I. Right: The distribution of H I masses of objects detected in AHISS (Zwaan 2000) and HIPASS (Meyer et al. 2004) are given in the upper and lower panel, respectively.

3.3.2 H I masses of the objects detected in WSRT CVn survey

Now that the model to estimate distances from the measured recession velocities has been adopted, we can proceed estimating the H I masses of 70 detections using the relation provided at the beginning of this section. There are 13 galaxies detected in the WSRT CVn survey with independent distance estimates available from the literature. These distances are given in Table 3.3 together with the name of the object, its position, the method used for the distance estimate and the reference to the measurement. For galaxies without an independent distance observation, recession velocities in the LG frame are used to obtain their distances. The distances are calculated as V_{LG}/H_0. We use $H_0 =$ 70 km s^{-1} Mpc^{-1}. Local Group velocities of all detections are given in Table 2.2 in Chapter 2. We discuss the effect of the uncertainties in the estimated distances on the H IMF in Section 3.4.5.

We present the distribution of $M_{\rm HI}$ masses assigned to the objects detected in the WSRT CVn survey in the left panel of Figure 3.3. The large histogram represents the distribution of H I masses of all 70 detections, while the filled histogram represents the distribution of 19 objects detected for the first time in the 21–cm emission line in the WSRT CVn survey. All new H I detections fall in the low-mass part of this distribution of H I masses. In the right panels of Figure 3.3, the distributions of H I masses constructed from AHISS (upper panel, Zwaan 2000) and HIPASS (lower panel, Meyer et al. 2004) are presented. The WSRT CVn survey peaks at a lower H I mass, $\sim 10^8$ $M_\odot$, when compared to AHISS and HIPASS, which peak at $\sim 10^9$ $M_\odot$ and $\sim 10^{9.75}$ $M_\odot$, respectively. From this comparison it is clear that these three blind H I surveys sample different volumes.

We use the minimum detectable integrated flux to estimate the maximum distance up to which an object with a certain HI mass can be detected in our survey. Using

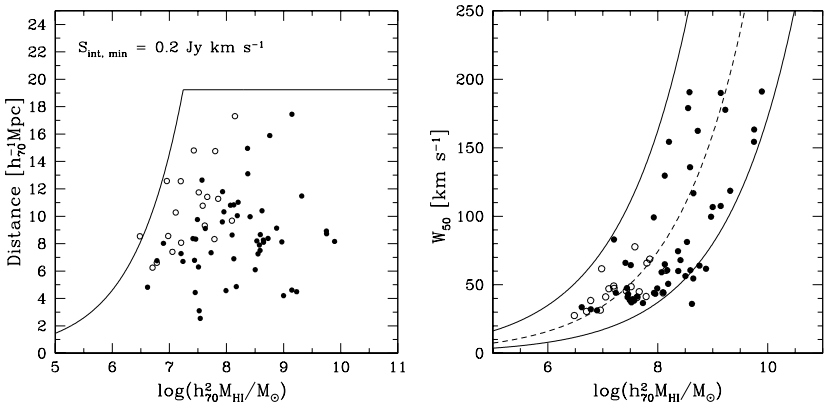

Figure 3.4: Left: Distances and H I masses of objects detected in the WSRT CVn survey. The continuous line represents the maximum distance up to which an object of a given H I mass can be detected in the WSRT CVn survey. Right: $W_{50} - M_{\rm HI}$ relation for the population of objects detected in the WSRT CVn survey. The dotted line corresponds to the relation $W_{50} = 0.16 M_{\rm HI}^{1/3}$ (Briggs & Rao 1993). The continuous lines mark the upper and lower limits of the given relation. In all panels, objects detected for the first time in the 21–cm emission line in the WSRT CVn survey are presented with empty symbols; otherwise they are presented with filled symbols.

real and simulated detections, we calculate that $S_{\rm min} = 0.2$ Jy km s^{-1} is the minimum detectable integrated flux to classify a detection in our survey as a real object. The minimum H I mass which an object must have to be detected in our survey is then given by $\log M_{\rm HI,min}/{\rm M}_\odot = 4.67 + 2 \log D$. Objects with H I masses 10^6 h$_{70}^{-2}$ $M_\odot$ can be detected up to a distance of 4.6 h$_{70}^{-1}$ Mpc and objects with H I masses 10^7 h$_{70}^{-2}$ $M_\odot$ up to 14.6 h$_{70}^{-1}$ Mpc. The distribution of H I masses of objects detected in the WSRT CVn survey as a function of their distance is presented in the left panel of Figure 3.4. The detection limit of the WSRT CVn survey as a function of H I mass of an object is indicated with a continuous line in the same panel. We adopt a value of 19.24 Mpc as the bandwidth depth D_{bw} of the survey. This distance is calculated from the maximum velocity covered in the datacubes produced from nights 10-60 of the observations. A correction of 17.8 km s^{-1} is added to this velocity, to account for the transformation from barycentric to LG velocity frame. The added value of 17.8 km s^{-1} is obtained by averaging the correction values used to transform velocities from the barycentric to the LG velocity frame for 70 WSRT CVn detections. The distance D_{bw} is obtained under assumption of a Hubble flow with $H_0 = 70$ km s^{-1} Mpc^{-1}.

In optically selected samples of H I detected galaxies, a correlation between the line width and the H I mass of an object has been seen (e.g. Briggs & Rao 1993). It can be described as $W_{50} \propto M_{\rm HI}^{\beta}$ and therefore it is an H I equivalent of the Tully-Fisher relationship (Minchin 2001). The relationship for the H I detections is based on the measured velocity widths, not corrected for inclination. The relationship found by Briggs & Rao (1993) is $W_{50} = 0.16 M_{\rm HI}^{1/3}$ and is presented as a dashed line in the right panel of Figure 3.4. The upper and lower limits of the relationship given by the same authors are presented with continuous lines in the same Figure. With some scatter, our detections

follow this relation.

3.4 The WSRT CVn H I mass function

3.4.1 Method

The luminosity function and H I MF are usually represented in the form of a Schechter (1976) function. The latter can be written as

$$\phi(M_{\rm HI}) = \frac{dN}{d\log(M_{\rm HI})} = \ln(10)\phi^* \left(\frac{M_{\rm HI}}{M^*_{\rm HI}}\right)^{\alpha+1} \exp\left(-\frac{M_{\rm HI}}{M^*_{\rm HI}}\right). \tag{3.3}$$

The parameter α is the slope at the faint end of the function, $M^*_{\rm HI}$ is the characteristic turn-over mass and ϕ^* is the normalisation factor.

A few methods have been developed to estimate the H I MF and the luminosity function. The $\Sigma(1/V_{max})$ method is most widely used for the determination of the H I MF. It has been developed by Schmidt (1968) and it is constructed by summing the weights given to the objects in a particular mass bin. The weight of each object is given by the reciprocal value of the maximum volume in which an object can be placed and still be detected in the survey. The uncertainty in each of the bins is obtained by summing the squared values of the same weights given to the objects in a particular mass bin.

The main shortcoming of the $\Sigma(1/V_{max})$ method is that it assumes that the population of objects used for the evaluation of the luminosity or mass function is distributed homogeneously in the observed volume. Even though the volume covered by the WSRT CVn survey is not homogeneous, we will refer to the results obtained by Zwaan et al. (1997) who concluded that the $\Sigma(1/V_{max})$ method used to determine the H I MF (based on 66 detections) is not very sensitive to the effects of the large structures. On the other hand, the maximum likelihood methods can produce erratic results in the estimation of the H I MF, when only few galaxies per bin are used (Zwaan et al. 1997).

3.4.2 Completeness

In Chapter 2 we investigated the completeness of the survey using synthetic sources inserted into the 3-dimensional line datacubes. The completeness of the survey was defined as the fraction of the sources recovered in the datacubes, using searching methods identical to the method applied for the real survey. We expressed the completeness of the survey as a function of the measured integrated flux $S_{\rm int}$ and peak flux in the integrated profile $S_{\rm peak}$ values of the sources.

Since there is a direct relation between the H I mass and $S_{\rm int}$, we determined the analytical form for the completeness of the survey only as a function of $S_{\rm int}$. For the fitting procedure we use the result obtained in Chapter 2. We consider the points, weighted according to the relative fraction of datacubes of a particular type in the real survey, as the best estimate of the completeness of the survey (see Chapter 2). Using the χ^2 minimisation procedure, the completeness of the survey can be described empirically in the form

$$C = erf(4.59[S_{\rm int}^{1/2} - 0.435]), \tag{3.4}$$

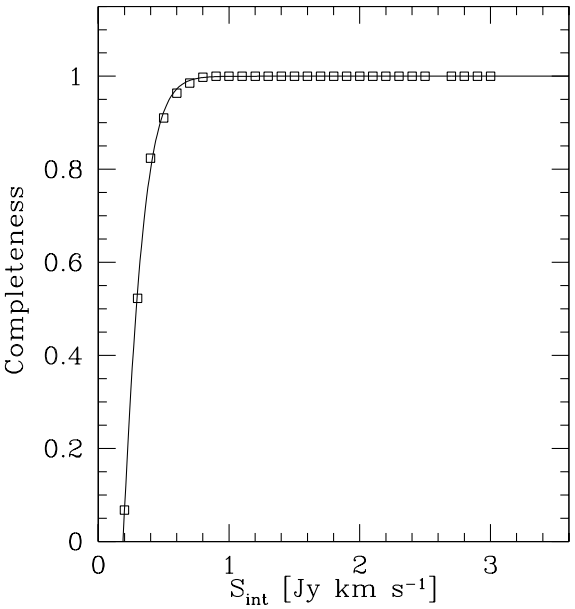

Figure 3.5: Completeness function of the WSRT CVn survey. The empty squares represent the fraction of objects detected in the simulations, weighted with the relative abundance of the type of datacube in which the detection resides. The continuous line is the analytic fit to the symbols obtained by χ^2 minimisation.

for S_{int} values expressed in Jy km s^{-1}. All points are weighted equally. We present this fit with a comparison to the binned completeness points in Figure 3.5.

The completeness function in its analytic form can be used to correct the number of detections with a certain S_{int} for those that have been missed by the survey, and can on the other hand be used to remove detections with a probability too low of being detected. Taking the completeness of the survey into account, the H IMF using the Schmidt method can be estimated weighting each detection with its maximum volume, corrected for the incompleteness of the survey $V_{max,C} = CV_{max}$.

3.4.3 Result

To estimate the H IMF using the Schmidt method, the data has to be binned. Given the size of the sample of H I objects detected in the WSRT CVn survey, a prerequisite to have enough points per bin and enough bins to make useful fits led to a choice of the bin size of 0.5 in logarithmic units of H I mass. Another choice has to be made for the minimum mass defining the starting bin. Of course, one would like to constrain the H IMF to as low H I mass as possible, but results obtained need to have significant reliability. We estimated the H IMF using different minimum H I masses in the starting bin, keeping the bin size fixed. The minimum H I mass at the centre of the starting bin is 6.5 in logarithmic units, chosen such that the first bin contains at least 2 detections. We repeat the process of re-binning the data and fit the Schechter function eight times, shifting the minimum H I mass in the starting bin by 0.05 in logarithmic units.

Figure 3.6 shows the result of the H IMF calculation for the WSRT CVn survey for the different minimum H I masses used in the binning. To fit the Schechter function, we bin the data in logarithmic intervals of half a decade and multiply the number of objects in each of the bins with 2 to scale our results to the commonly used H IMF representation in

decade-wide bins. The open circles correspond to the H IMF calculated using all available objects and weighting them with $V_{\rm max} = 1/3\Omega D^3_{\rm max}$. The parameter Ω is the area covered by the survey and $D_{\rm max}$ is the maximum distance up to which an object with a given H I mass can be detected in the survey. $D_{\rm max}$ can be calculated using the minimum integrated flux detectable by the WSRT CVn survey of 0.2 Jy km s^{-1}. For objects for which $D_{\rm max}$ is larger than the adopted bandwidth of the survey $D_{\rm bw}$ = 19.24 Mpc, the latter is used as the maximum distance up to which such an object can be detected. The solid circles represent the H IMF from the binned data points, weighted with a maximum volume corrected for the incompleteness of the survey $V_{\rm max,C} = CV_{\rm max}$. For the H IMF data points which are identical in both estimates of the H IMF, solid symbols overlay the empty ones. There is one object which is statistically not-detectable in this survey (object WSRT-CVn-43). The completeness for this object is smaller than zero, and therefore, this object is excluded from the evaluation of the H IMF corrected for the incompleteness of the survey to detect objects with small $S_{\rm int}$ values. We fit the Schechter function only to the binned data corrected for incompleteness, using the χ^2 fitting. The errors in each bin are calculated using the summation $\Sigma(1/V_{max,C})$.

The best-fit parameters, calculated for different minimum H I masses in the first bin, together with a goodness–of–fit and estimated errors in the fitted parameters, are given in Table 3.4. The errors correspond to the 1σ confidence intervals from the χ^2 minimisation obtained individually for each of the parameters. The average values of the best–fit parameters are given in Table 3.5. Given that the Schechter function is a combination of a power law and an exponential function which has to be fitted to the binned data points simultaneously, and that there are only a few detections which constrain the high-mass end of the H IMF, we compare only the slope α of the different fits. We chose the fit obtained with the starting bin centred at 6.65 of the H I mass in logarithmic units as the best representation of the data observed in the WSRT CVn survey. The slope obtained with this binning has the smallest difference from the average slope, both calculated as the arithmetic average and calculated as the average value weighted with the minimum χ^2 values. In Figure 3.7 we present again the H IMF fit to the binned data starting with a bin centred on a logarithmic H I mass of 6.65 together with the two Schechter functions for the average parameters as summarised in Table 3.5. The adopted best-fit Schechter function of the WSRT CVn blind H I survey has parameters $\alpha = -1.17$, $\log(h^2_{70} M^*_{\rm HI}/M_\odot) = 9.57$ and $\phi^* = 0.125 h^3_{70}$ Mpc^{-3} dex^{-1}.

3.4.4 Low number statistics

Unlike in most H I surveys carried out up to date, which suffer from low-number statistics at the low–mass end of the H IMF, the uncertainties in the H IMF points estimated from the WSRT CVn survey are larger at the high–mass end. A detailed discussion on the differences between the surveys will be presented in Subsection 3.6.1. The number of H I detections which constrain the $M^*_{\rm HI}$ parameter in the WSRT CVn survey is very small. There are only 3 detections with a mass above $M^*_{\rm HI}$ of the adopted best–fit Schechter function.

To test the stability of the resulting slope obtained from this survey, we calculate the H IMF of the WSRT CVn survey adopting $M^*_{\rm HI} = 7.24\times\ 10^9$ h$^{-2}_{70}$ $M_\odot$ obtained from the HIPASS (Zwaan et al. 2005b). The best-fit values of the remaining two Schechter parameters from the WSRT CVn survey data are α = -1.22 and $\phi^* = 0.085 h^3_{70}$ Mpc^{-3}

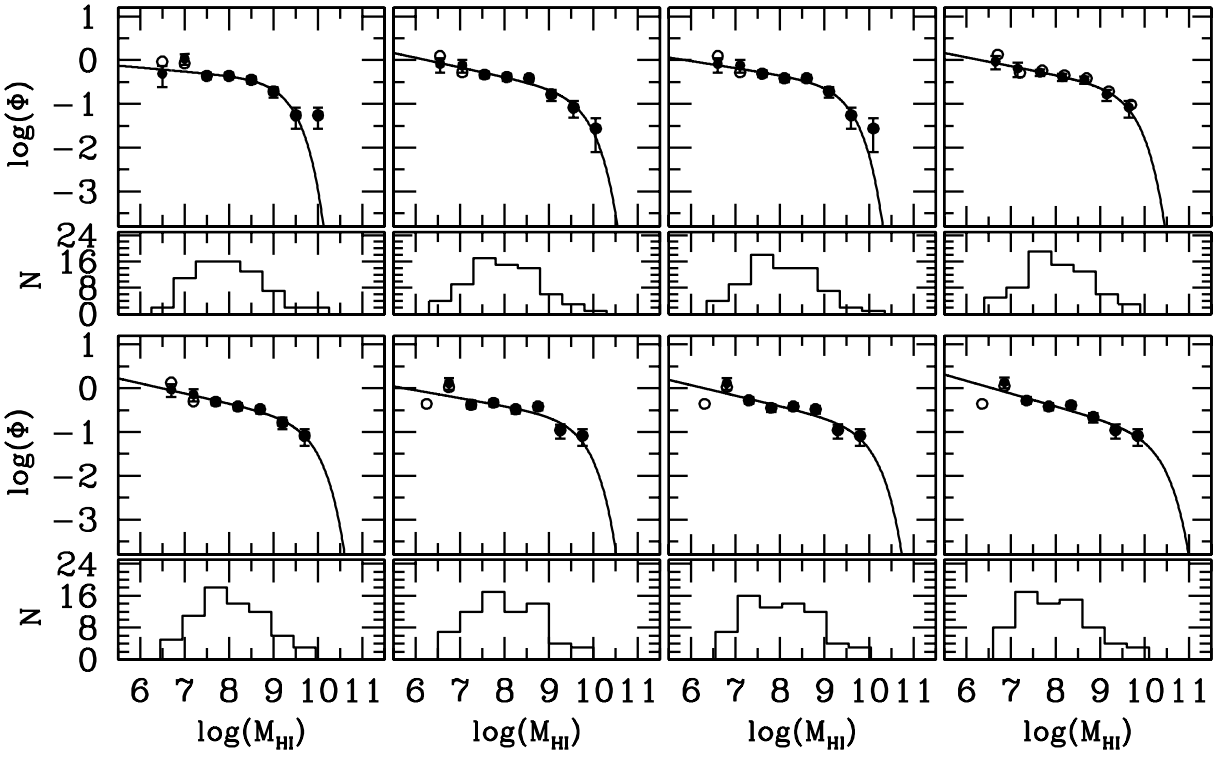

Figure 3.6: HIMF dependence on the binning. In the upper part of each panel the empty symbols represent the HIMF obtained by weighting each detection with V_{max} (70 detections) and the solid symbols represent the HIMF obtained by weighting each detection with $V_{max,C}$ (69 detections). For bins which contain only galaxies for which the survey is complete, the solid symbols overlay the open ones. Only errors in the HIMF points corrected for incompleteness are presented. The lines indicate the best-fit Schechter functions to the completeness-corrected HIMF. Histograms of the HI masses (69 detections) with the first bin centred at the same mass as the corresponding HIMF, are presented in the lower part of each panel. The centres of the bin with the minimum mass expressed in logarithmic units of $M_\odot$ are: 6.5, 6.55, 6.6, 6.65, 6.7, 6.75. 6.8 and 6.85, going from left to right, first upper than lower panels. $M_{\rm HI}$ is in units of $h_{70}^{-2} M_\odot$, ϕ is in units of h_{70}^3 Mpc^{-3} dex^{-1}.

$M_{HI,c}$	χ^2	α	$\sigma\alpha$	$\log(M^*_{\rm HI}h^2_{70}/M_\odot)$	$\sigma\ \log(M_{\rm HI}h^2_{70}/M_\odot)$	Φ^*	$\sigma\Phi^*$
-	-	-	-	-	-	[h^3_{70} Mpc^{-3} dex^{-1}]	[h^3_{70} Mpc^{-3} dex^{-1}]
6.50	7.36	-1.07	-0.98	9.22	9.05	0.187	0.130
			-1.14		9.42		0.257
6.55	2.93	-1.18	-1.10	9.62	9.32	0.107	0.067
			-1.25		9.91		0.169
6.60	2.79	-1.13	-1.03	9.41	9.23	0.152	0.095
			-1.21		9.67		0.217
6.65	1.18	-1.17	-1.10	9.57	9.35	0.125	0.078
			-1.24		9.87		0.181
6.70	0.61	-1.20	-1.12	9.71	9.47	0.102	0.06
			-1.27		10.08		0.154
6.75	10.24	-1.15	-1.05	9.63	9.31	0.111	0.063
			-1.23		10.03		0.183
6.80	6.88	-1.21	-1.13	9.85	9.53	0.008	0.033
			-1.29		10.60		0.131
6.85	5.10	-1.27	-1.20	10.14	9.70	0.051	0.01
			-1.34		12.00		0.09

Table 3.4: Parameters of the H I IMF fits and their errors for different binning. The centre of the first non-empty bin is given in the first column.

	α	$\log(M^*_{\rm HI}\ {\rm h}^2_{70}/{\rm M}_\odot)$	ϕ^*
	-	-	$[{\rm h}^3_{70}\ {\rm Mpc}^{-3}\ {\rm dex}^{-1}]$
arithmetic average	-1.17	9.72	0.114
average weighted with χ^2	-1.18	9.69	0.112

Table 3.5: Average values of the H IMF parameters obtained by changing the minimum H I mass in the first non-empty bin.

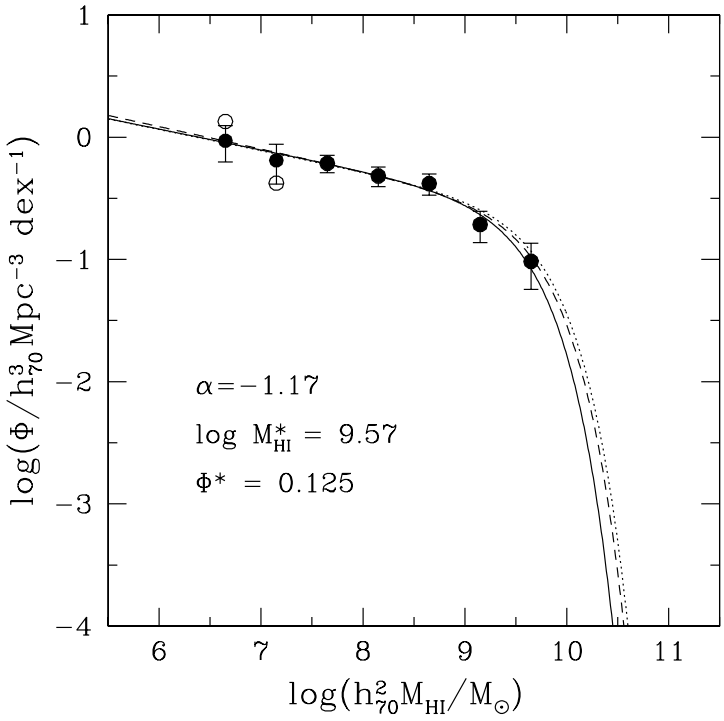

Figure 3.7: The WSRT CVn survey H IMF. The empty circles correspond to the H IMF calculated weighting the objects with $V_{max} = 1/3\Omega D^3_{\rm max}$. The solid circles are H IMF from the binned data points weighted with a maximum volume corrected for the incompleteness of the survey $V_{max,C} = CV_{max}$. For bins which contain only galaxies for which the survey is complete, the solid symbols overlay the open ones. The first data bin is centred on an H I mass of 6.65 in logarithmic units. The adopted best-fit Schechter function to these points is presented with a continuous line. Schechter functions constructed from the average parameters, obtained from the χ^2-minimisation to the 8 offsets binning, are presented with a dotted line (the arithmetic average of parameters) and a dashed line (the χ^2-weighted average).

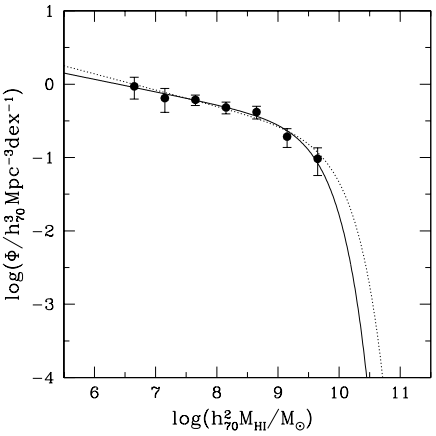

Figure 3.8: H IMF from the WSRT CVn survey using the $M^*_{\rm HI}$ parameter obtained from HIPASS (the dotted line). For comparison, the adopted H IMF from the WSRT CVn survey is also presented (the continuous line). The steepening of the WSRT CVn slope caused by adopting $M^*_{\rm HI}$ from HIPASS ranges from -0.01 to -0.1 within the 1σ confidence intervals of the fitted slope.

dex-1 with $\chi^2 = 2.10$. The lowest α value within the 1σ confidence interval is -1.27, and the highest value is -1.18. For ϕ^*, the 1σ uncertainties range from 0.068 to 0.100 h^3_{70} Mpc^{-3} dex-1. Adopting the $M^*_{\rm HI}$ parameter from HIPASS, obtained from statistics based on a few hundred of galaxies with $M_{\rm HI} > M^*_{\rm HI}$, will cause a steepening of the slope obtained from the WSRT CVn survey by an amount of -0.01 to -0.1 within the 1σ confidence intervals of the new slope.

3.4.5 Uncertainties in the H I parameters

The H I parameters of the objects in the WSRT CVn survey have some uncertainties. We already discussed the effect of different choices for the precise binning. Here, we will explore the influence of the uncertainties in the measured distances and integrated fluxes of the detections on the Schechter parameters of the H IMF.

In Subsection 3.3.1 we explored the different approaches to model the velocity field in the CVn volume in order to estimate distances to objects detected in the WSRT CVn survey with the highest possible accuracy. Our conclusion was to use the LG recession velocities to obtain distances to the detections, assuming a Hubble flow with $H_0 = 70$ km s^{-1} Mpc^{-1}. Our analysis was based on 64 galaxies in the CVn constellation area with distances estimated independently of their recession velocity. This sample of galaxies is distributed around the Hubble flow with $H_0 = 70$ km s^{-1} Mpc^{-1} with a rms dispersion of approximately 178 km s^{-1}.

In the following analysis, we adopt the same peculiar velocities for the WSRT CVn galaxies as observed in the larger area of the CVn constellation. All detections in the WSRT CVn survey have recession velocities below 1230 km s^{-1} and a peculiar velocity of about 180 km s^{-1} (the 1σ uncertainty from the sample of 64 galaxies in this region) will cause an error on the estimated distance of 15% for a detection at $\sim$1200 km s^{-1} and an error of 45% for a detection at $\sim$400 km s^{-1}. When assuming more realistic peculiar velocities of about 290 km s^{-1} at at $\sim$1200 km s^{-1} and of about 100 km s^{-1} at $\sim$400 km s^{-1} (see Table 3.2), distances of galaxies at these two recession velocities will have a typical error of 25%.

To test whether the uncertainties in distances of the detections in the the WSRT CVn

sample affect the parameters of the best-fit Schechter function, we set up simulations adding assumed peculiar velocities to the measured recession velocities of the detections from the WSRT CVn survey. A first set of simulations (called simvpec1 from here on) consists of drawing a peculiar velocity from a Gaussian distribution with a dispersion of 178 km s^{-1} and adding this value to the recession velocity of each of the detections, except for those detections for which there exists a literature value for the distance (from a primary or secondary distance indicator). For the latter galaxies, we add a value to the distance, randomly chosen to correspond to an error of 10% or 25%, depending on the type of the distance indicator. In the second set of simulations (simvpec2), we use equation 3.2 to estimate the width of a Gaussian distribution from which the value of a peculiar velocity is selected randomly for detections with recession velocities less or equal to 700 km s^{-1}. For the rest of the detections, we draw a value of a peculiar velocity from a Gaussian distribution with a dispersion of 290 km s^{-1}. This value is equal to the dispersion of differences between recession velocities larger than 700 km s^{-1} and the velocities derived from the Hubble flow with $H_0 = 70$ km s^{-1} Mpc^{-1} of the sources studied in Subsection 3.3.1. Uncertainties in the distances obtained using independent distance indicators have been calculated in the same manner as in the simvpec1 set of simulations. Each set of simulations consists of 100 runs. The number of simulations is sufficient to achieve convergence of the dispersion in the parameters obtained from the simulations.

The results from the simulations are presented in Figure 3.9. In the left panel, the adopted best–fit Schechter function from the WSRT CVn survey is presented with a continuous line. The Schechter functions, constructed from the χ^2–weighted average parameters obtained from 100 simulations each, are presented with a dotted line (simvpec1) and a dashed line (simvpec2). The errors are presented in the right panel in Figure 3.9. For the WSRT CVn H IMF 1σ confidence intervals obtained from the χ^2-minimisation are plotted. The errors in the estimated parameters from the simulations – both arithmetic averages and χ^2–weighted averages – have been calculated as the standard deviations of the 100 simulations. It is obvious that the slope of the WSRT CVn H IMF is not affected by the modelled uncertainties of the distances. All α parameters obtained from the simulations fall in the area limited by the 1σ contour of the best–fit slope of the WSRT CVn H IMF. All ϕ^* parameters from the simulations and $M^*_{\rm HI}$'s from simvpec2 agree within the 1σ errors of the corresponding parameters from the WSRT CVn survey. The M^*_{HI} parameter calculated as the χ^2–weighted average in simvpec1 has been affected the most: it is about 2.4 times larger than the best–fit WSRT CVn $M^*_{\rm HI}$ value.

Apart from the distances, $S_{\rm int}$ is the second parameter which figures in the $M_{\rm HI}$ calculation. This parameter determines the correction for the incompleteness of the survey, which has been applied during the H IMF calculation to each of the objects. The $S_{\rm int}$ values of the objects detected in the WSRT CVn survey have been calculated by averaging the $S_{\rm int,c}$ values obtained by integrating the pixels in a mask defined around a detection, and the $S_{\rm int,MB}$ values obtained by integrating the flux in a box around a detection. In Chapter 2, we estimated the uncertainty in the integrated fluxes using the synthetic sources. The uncertainties are 0.240 Jy km s^{-1}, 0.421 Jy km s^{-1} and 0.242 Jy km s^{-1} for the $S_{\rm int}$, $S_{\rm int,c}$ and $S_{\rm int,MB}$ values, respectively. The uncertainties are small, but there are 13 objects with $S_{\rm int} \leq 1$ Jy km s^{-1} and for those objects the estimated uncertainties of $S_{\rm int}$ produce errors ranging from 24% to 136%.

Given that the estimation of the slope of the H IMF at very low H I masses is one

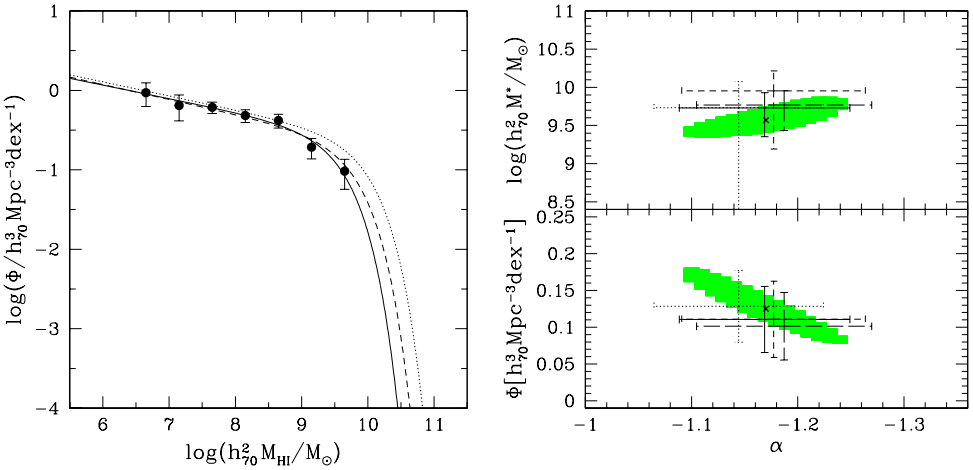

Figure 3.9: Left: The H IMFs modelling the uncertainties in distances. The best–fit H IMF from the WSRT CVn survey is presented with a continuous line. The H IMF obtained from the χ^2–weighted average values from 100 simulations in the simvpec1 set is presented with a dotted line. The H IMF obtained from the χ^2–weighted average values from 100 simulations in the simvpec2 set is presented with a dashed line. Right: Uncertainties of the H IMF parameters due to the distance uncertainties. Filled areas correspond to 1σ confidence intervals of the best-fit parameters of the H IMF from the WSRT CVn survey. The best-fit parameters of the Schechter function from the WSRT CVn survey are marked with a cross. Standard deviation intervals of the best-fit parameters obtained from 100 simulations are marked with a dotted line for the parameters obtained by averaging the values from simulations simvpec1, a short-dashed line for the χ^2–weighted average values from simulation simvpec1, a continuous line for the average values from simvpec2 simulations and long–dashed line for the χ^2–weighted average values from the simvpec2 simulation.

of the main goals of the WSRT CVn survey, it is of great importance to investigate the effect of the large relative errors in the smallest integrated fluxes. We investigate the influence of the integrated flux uncertainties setting up two sets of simulations. A first set of simulations is based on adding an additional amount of integrated flux to the S_{int} value measured (simsflux1). This additional amount of integrated flux has been drawn from a Gaussian distribution with a dispersion of 0.240 Jy km s^{-1}. The second set of simulations consists of changing the $S_{\rm int,c}$ and $S_{\rm int,MB}$ values independently by a value randomly chosen from a Gaussian distribution with a dispersion of 0.421 Jy km s^{-1} and 0.242 Jy km s^{-1} for $S_{\rm int,c}$ and $S_{\rm int,MB}$ parameters (simsflux2). We made an approximation in the simulation setups, assuming that uncertainties in integrated fluxes follow a Gaussian distribution. For objects with WSRT-CVn indexes 67A and 67B we used the same uncertainty as for the rest of the objects, even though the values of their integrated fluxes have been taken from the literature. In total, we carried out 100 simulations for each of the cases explored, sufficient to ensure convergence of the dispersions in the Schechter parameters fitted to each of the simulation run.

The resulting H IMFs and the average values of the Schechter parameters from the simulations are presented in Figure 3.10 in the left and the right panels, respectively. The

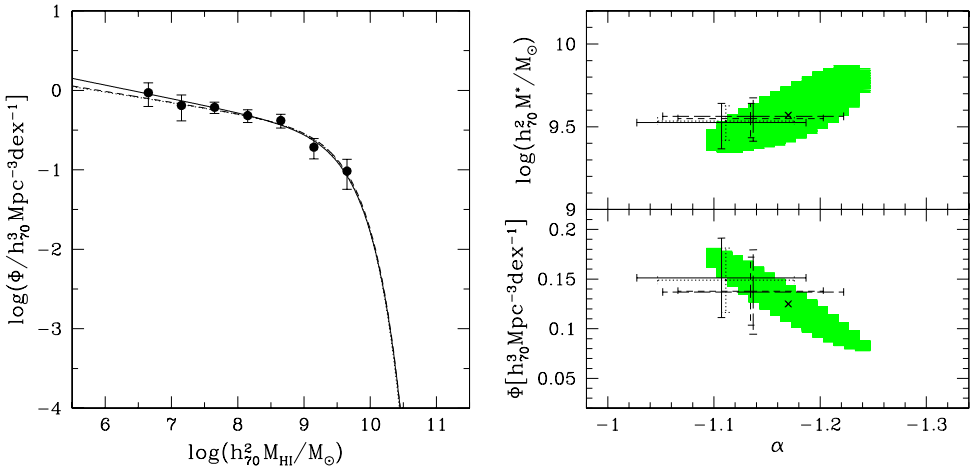

Figure 3.10: Left: The H IMFs from modelling the uncertainties in integrated fluxes. The best–fit H IMF from the WSRT CVn survey is presented with a continuous line. The H IMF obtained from the χ^2–weighted average values from 100 simulations in the simsflux1 set is presented with a dotted line. The H IMF obtained from the χ^2–weighted average values from 100 simulations in the simsflux2 set is presented with a dashed line. Right: Uncertainties in the H IMF parameters due to uncertainties in integrated flux. Filled areas correspond to 1σ confidence intervals of the best-fit parameters of the H IMF from the WSRT CVn survey. The best-fit parameters of the Schechter function from the WSRT CVn survey are marked with a cross. Standard deviation intervals of the best-fit parameters obtained from 100 simulations are marked with a dotted line for the parameters obtained by averaging the 100 values from the simsflux1 simulations, a short-dashed line for the χ^2–weighted average values from the simsflux1 simulations, a continuous line for the average values from the simsflux2 simulations and long–dashed line for the χ^2–weighted average values from the simsflux2 simulations.

simulated H IMFs are practically indistinguishable at the high–mass end, while there is a slight flattening of the slopes obtained from the simulations compared to the best–fit α from the WSRT CVn survey. The slope flattening $\Delta\alpha$ is 0.059 and 0.036 for α values obtained from the arithmetic averaging of parameters from the 100 simsflux1 simulations, and from the χ^2–weighted averaging of 100 best–fit α values from 100 simsflux1 simulations, respectively. Compared to the slope in the WSRT CVn survey, the 100 simflux2 simulations give a slope which is flatter by 0.063 for the arithmetic average and flatter by 0.033 for the χ^2–weighted average.

3.5 Comparison with previous H I MF estimates

It is of interest to compare the H IMF obtained from the WSRT CVn survey with the H IMF estimates from other recent H I surveys. In Figure 3.11 we reproduce the adopted best-fit parameters of the WSRT CVn Schechter function with a continuous line. The function has been re-normalised to match the number–density at the high–mass end of the H IMF from the recent major blind H I surveys as presented in Figure 3.11. We

find that the volume probed by the WSRT CVn survey is over-dense 11 times with respect to the volumes covered by the previous surveys. The best-fit Schechter function obtained from the HIPASS is presented with a short–dashed line (Zwaan et al. 2005b), the H I MF fit from the Arecibo Dual Beam Survey (ADBS) is presented with a long–dashed line (Rosenberg & Schneider 2002) and the resulting H I MF obtained from the AHISS is presented with a dot–dashed line (Zwaan et al. 1997). We plot all the curves converted to $H_0 = 70$ km s^{-1} Mpc^{-1}. In Figure 3.12 we present the errors in the α and $M^*_{\rm HI}$ parameters from the WSRT CVn survey and the three additional major blind H I surveys discussed above. For a comparison, we include α and $M^*_{\rm HI}$ values and their errors from the H I MF estimated from the optically selected H I sample presented by Springob et al. (2005) in Figure 3.12 as well.

The three H I MF functions from the major blind H I surveys agree quite well at the high–mass end. The $M^*_{\rm HI}$ value obtained in this survey is the lowest, and agrees with the $M^*_{\rm HI}$ estimates from AHISS and HIPASS within the 1σ errors. The $M^*_{\rm HI}$ parameter of the H I MF of the optically selected H I detections has a larger value, which is slightly outside of the 1σ confidence interval of $M^*_{\rm HI}$ from the WSRT CVn survey. It is obvious that the uncertainty in the $M^*_{\rm HI}$ value is largest in the WSRT CVn survey compared to the uncertainty in this parameter as derived from the other surveys considered. This is not surprising given the small number of galaxies at the high–mass end of the WSRT CVn H I MF. The absence of massive galaxies in the CVn groups of galaxies is real (e.g. Karachentsev et al. 2003b) – it can not be caused by any presumed inability of the WSRT CVn survey and the detection technique used to detect objects with high H I masses.

The situation is completely different at the low–mass end of the H I MF. The α parameter of the functions presented span a range of values between -1.17 and -1.53. The slope at the low–mass end of the H I MF obtained from the WSRT CVn survey agrees very well with the slopes obtained from the AHISS HIMF and the HIMF based on the optically selected galaxies in Springob et al. (2005). The predicted number of H I objects with small masses with respect to the number of high–mass H I objects is a few times smaller than the corresponding predicted ratio from the HIPASS, and especially from the ADBS. As can be seen from Figure 3.12, the uncertainty in the α value estimated from the WSRT CVn survey is smaller than the uncertainties in α as obtained from the AHISS and the H I MF of optically selected galaxies in Springob et al. (2005). However, it is comparable to the error from the ADBS and it is larger than the error in the slope fitted to the HIPASS data.

3.6 Discussion

The discrepancies at the low–mass end of the H I MF can have different reasons. Some of them can be caused by the differences in the surveys themselves, as well as differences in how the sensitivity and completeness of the surveys have been treated. The small number statistics, and especially the small number statistics at the low–mass end of the H I MF, present in the majority of the H I surveys also plays an important role. The low–number statistics in the H I surveys is closely related to the limitations imposed by the existing radio telescopes, which can detect objects with H I masses less than 10^8 $M_\odot$ up to a distance of only a few tens of Mpc. At such small distances, uncertainties in distances will be large and they may affect the estimated H I masses and therefore the

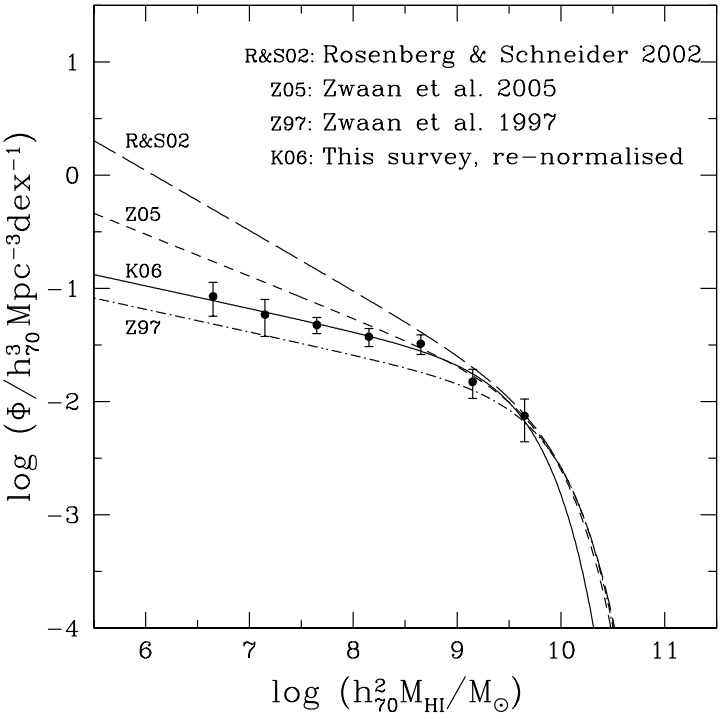

Figure 3.11: Comparison of the H IMFs estimated from recent blind H I surveys.

estimated slope of the H IMF. The astrophysical characteristics of the volumes probed by the H I surveys may also play a role in the broad range of α values. The H IMFs of different morphological types are characterised with rather different slopes (Zwaan et al. 2003; Springob et al. 2005), and in environments of different densities the slope of the H IMF may be influenced by the overall density of the volume surveyed (Springob et al. 2005; Zwaan et al. 2005b).

3.6.1 Specifications of the surveys

The H IMFs from the literature with which we compared the WSRT CVn H IMF are derived from the H I surveys which can be considered as representative. In Chapter 1 we presented an overview of recently and currently carried out blind H I surveys (Table 1.1. This overview includes also the blind H I surveys for which the H IMF derived has been presented here in Figure 3.11. AHISS (Zwaan et al. 1997) and ADBS (Rosenberg & Schneider 2002) are both drift–scan surveys carried out with the Arecibo telescope. HIPASS (Barnes et al. 2001) has been carried out with the Parkes 64 m telescope, using the 21–cm multi-beam receiver (Staveley-Smith et al. 1996). Statistics of the surveys are based on 66, 265 and 4315 detections in AHISS, ADBS and HIPASS, respectively. The total number of objects with small H I masses is very small: 1 detection with $M_{\rm HI} < 10^8$ h_{70}^{-2} $M_\odot$ in AHISS, 7 detections with $M_{\rm HI} < 10^8$ h_{75}^{-2} $M_\odot$ in ADBS and 41 detections

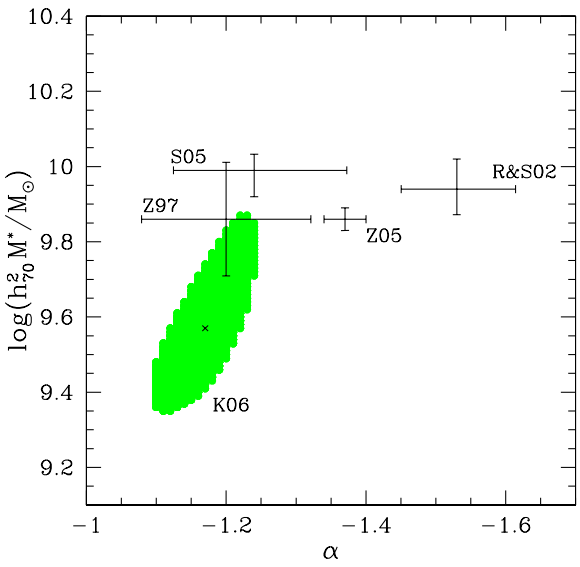

Figure 3.12: Comparison of the confidence intervals of the H IMFs estimated from recent blind H I surveys. We use the same notification as in Figure 3.11. In addition, S05 stands for the H IMF derived from the optically selected galaxies by Springob et al. (2005).

with $M_{\rm HI} < 10^8$ h$_{70}^{-2}$ $M_\odot$ in HIPASS. For the H IMF calculation based on the HIPASS data, only detections with a completeness larger than 0.5 have been used, which limits the lowest H I mass detection used for the H IMF in this survey to be $10^{7.2}$ h$_{75}^{-2}$ Mpc. The searching process in AHISS and ADBS relies on by-eye examination of the H I spectra. A HIPASS catalogue of detections has been produced by running a semi-automatic searching algorithm, based on the integrated peak flux maximum cut-off, and additional by-eye examination and possible elimination of detections revealed by the searching algorithm. Completeness of the AHISS was addressed in an analytical manner, using the $W_{50} - M_{\rm HI}$ relation from Briggs & Rao (1993), and in the ADBS and HIPASS using the recovery fraction of synthetic sources with a range of H I properties inserted in the H I data. AHISS is the most sensitive survey, with a 1σ rms noise of $\sim$0.75 mJy Beam^{-1} per 32 km s^{-1}. The 1σ rms value in the ADBS is $\sim$3.5 mJy Beam^{-1} per 32 km s^{-1} and in the HIPASS this is $\sim$13 mJy Beam^{-1} per 18 km s^{-1}.

The WSRT CVn survey is one of the very few blind H I surveys carried out using a synthesis telescope. The searching process has been carried out in a completely automated manner. The completeness of the survey has been derived from the recovery fraction of synthetic sources. The completeness of the survey has been described as a function of two parameters, $S_{\rm int}$ and $S_{\rm peak}$. Given that the majority of detections in the WSRT CVn survey are small and of simple structure, with small H I fluxes, narrow profile widths of a Gaussian shape and small in (H I) size, we do not expect that we make an substantial error in addressing the completeness of the survey as a function of one parameter only. When considering the properties of the H I data, the WSRT CVn survey is of similar sensitivity as the AHISS survey with a mean 1σ rms noise of 0.837 mJy Beam^{-1} per $\sim$33 km s^{-1}.

All surveys discussed have their own specifications, but there is no obvious hint which would lead to an understanding of the differences observed in the slopes of the H IMF.

The surveys discussed here have the largest number of detections of all independent blind H I surveys completed up to date. Most of the values for the slope α for the H IMFs derived from yet other surveys fall in the interval from -1.17 to -1.53 established from the surveys discussed in the detail here. One of the greatest puzzles of the 21-cm HIMFs derived to date is the result obtained by Schneider et al. (1998), who claimed a sharp steepening of the H IMF below H I masses of 10^8 $M_\odot$. On the other hand, Verheijen et al. (2001) reported a nearly flat H IMF, derived from a blind H I survey in the Ursa Major cluster which covered about 16% of the cluster volume.

For us, the result obtained by Kraan-Korteweg et al. (1999) is of great interest. Kraan-Korteweg et al. (1999) have carried out a driftscan survey in the 21–cm emission line using the Nançay radio telescope in the direction of the CVn constellation. The survey overlays partially with ours: a strip ranging from 11^h 30^m to 15^h 00^m in right ascension and from $+29^0$ $08^{'}$ to $+35^0$ $22^{'}$ in declination has been surveyed. It covers the CVn groups of galaxies as well as the ComaI cluster and a void region. The survey is sensitive to detect objects with an H I mass of 1-2 $\times$ 10^8 h_{100}^{-2} $M_\odot$ at 23 h_{100}^{-1} Mpc and 4-8 $\times$ 10^7 h_{100}^{-2} $M_\odot$ throughout the CVn groups, where the exact mass value depends on the position of the detection with regard to the centre of the telescope beam. There are 33 reliably detected galaxies and the H IMF constructed has a slope of -1.20, which is in excellent agreement with our result.

In comparing the Schechter parameters and their confidence intervals obtained from the blind H I surveys presented in Figure 3.12, we also included the parameters and their errors from the optically selected H I sample of galaxies obtained by Springob et al. (2005). Their slope has a value of -1.24, in agreement with the slope of $\sim$-1.25 as estimated from the only other H IMF based on a large sample of optically selected H I detections which covers a range of morphological types from Sa to Irr, as presented by Briggs & Rao (1993).

3.6.2 The influence of distance uncertainties

Rosenberg & Schneider (2002) used the velocity flow model by (Tonry et al. 2000) to assign distances to most of the detected H I sources in the ADBS. In order to investigate the influence of the distance uncertainties on the H IMF, they added random Gaussian noise to the distances of synthetic sources created to resemble detections in the ADBS. Rosenberg & Schneider (2002) concluded that there is no significant change in the slope of the ADBS H IMF due to distance uncertainties, if the assumed velocity dispersion is less than 600 km s^{-1}. In a similar manner, Zwaan et al. (2003) added Gaussian noise to the distances of simulated sources, mimicking to the 1000 brightest detections (Brightest Galaxy Catalogue, BGC) in the HIPASS with distances estimated from LG recession velocities. A realistic velocity dispersion of 50-100 km s^{-1} would cause a steepening of the H IMF by no more than 0.05. The Mould et al. (2000) parametric distance model would change the best-fit slope of the H IMF, derived from the complete HIPASS catalogue, from -1.37 to -1.39. Springob et al. (2005) used the Tonry et al. (2000) velocity model to obtain distances to the majority of optically selected H I detections.

Masters et al. (2004) conducted a study to investigate the influence of the distance uncertainties on the slope of the luminosity and H I mass functions in the Local Universe. As Masters et al. (2004) in their conclusions point out, the use of a *good* model for the local velocity field is essential for deriving an unbiased H IMF. For the WSRT CVn survey

we first showed in Subsection 3.3.1 that the pure Hubble flow is a good approximation of the velocity field in the CVn area (if not for the Local Universe). After having justified the usage of the LG recession velocities to derive distances to our detections, we showed in Subsection 3.4.5 the effect of the velocity dispersion observed on the final Schechter function. All new slopes fell within the 1σ confidence interval of the best-fit WSRT CVn α value.

3.6.3 Dependence on morphology

In order to investigate whether the H I MF depends on morphological type of galaxies, numerous detections are needed. Almost three decades after the first heroic blind H I survey has been carried out, with a score of 1 detection (Shostak 1977), HIPASS finally allows studies of the H I MF dependence on morphological type. The BGC has been used for this purpose. Morphological classification of the optical counterparts of the H I detected sources in the BGC is conducted using HYPERLEDA. Based on this morphological classification, Zwaan et al. (2003) detect a steepening of the slope of the H I MF from $\alpha \sim$ -1.0 derived for all three samples with only Sa-Sb, Sbc-Sc or Scd-Sd galaxies to $\alpha \sim$ -1.4 fitted to the sample composed from Sm-Irr types of galaxies. Their finding agrees within 1σ uncertainties with the H I MF dependence on the morphological type of the optically selected H I detections in Springob et al. (2005) for Sa-Sb, Sbc-Sc and Sm-Irr types. For the subsample of Scd-Sd galaxies, Springob et al. (2005) found a slope of -1.44. The morphological classification has been conducted primarily using UGC types. The difference in the slopes of the H I MF for the Scd-Sd types of galaxies could be caused by differences in the morphological classification.

The number of detections in the WSRT CVn survey is to small to investigate the dependence of the H I MF on the morphologies of galaxies specifically. The sample is dominated by late-type galaxies, where about half of the detections with the available morphological classification are classified to be of Sm-Irr type (see Table 2.2 in Chapter 2).

3.6.4 Dependence on the environment

Recently, a lot of effort has been made to explain the effect of the environment on the properties of galaxies and their distributions. The H I surveys are much smaller then the optical surveys, and the only way to study the environmental effects is a direct comparison of the estimated H I MF from the surveys carried out in different environments (clusters, groups, voids). The main conclusion from this comparison is that H I galaxies detected in dense environments have an H I MF characterised by a flatter slope compared to the H I MF derived for the field (e.g. Springob et al. 2005, and references therein).

Springob et al. (2005) investigated for optically selected H I galaxies not only the dependence of the H I MF on morphological type, but also on the environment. Their subsamples of galaxies in environments of different densities contain similar fractions of the morphological types. In that case the differences in the derived H I MFs can be easily linked to the density of the environment. The local matter density has been derived from the reconstructed mass densities based on the IRAS Point Source Catalog Redshift Survey (PSCz, reconstruction presented in Branchini et al. 1999). Although the results obtained had statistically small differences, the conclusion was that the H I MF in the higher-density region shows a flattening at the faint end and lower values of $M^*_{\rm HI}$. In the

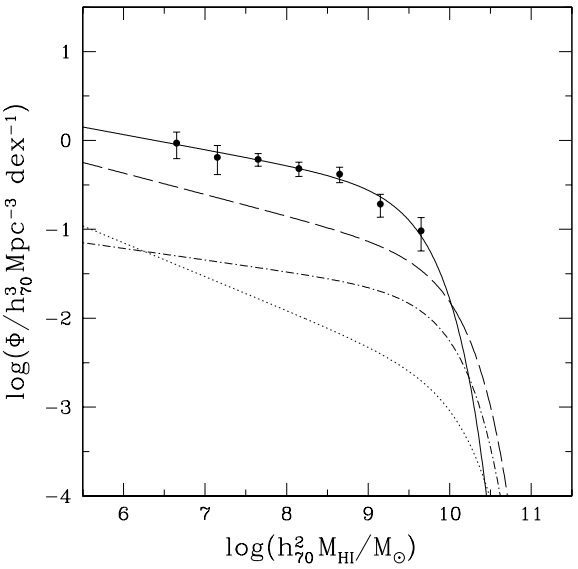

Figure 3.13: H IMF dependence on the environment. The H IMF derived in the volumes of different densities are presented with: dotted line for $n < 1.5$, dot-dashed line for $1.5 < n < 3.0$ and long–dashed line for density $n > 3$. Densities are given in the units of the local matter density. The WSRT CVn H IMF, corrected for incompleteness, is presented with solid circles and its best-fit Schechter function is presented with a continuous line.

terms of densities used by Springob et al. (2005), the volume probed by the WSRT CVn survey would be about six times denser than their densest region (by comparing the ϕ^* values).

Zwaan et al. (2005b) used the complete HIPASS catalogue to study the dependence of the H IMF on environment. The observed trend of the low-mass slope is opposite to the one found by Springob et al. (2005). The local density has been derived from counts of the nearest H I neighbours. Considering the different methods used to quantify the density of the environment in the studies by Zwaan et al. (2005b) and Springob et al. (2005), and the low statistical significance of the results by the latter

3.7 Conclusions

We have used the 70 H I detections and their parameters detected in the WSRT CVn blind 21–cm emission line survey to construct the H IMF. The sample covers a range of H I masses from 6.48 to 9.89 expressed in logarithmic units of $h_{70}^{-2}\ M_\odot$. A fraction of 37% of the detected objects has H I masses below $10^8\ h_{70}^{-2}\ M_\odot$, making this the H I selected sample with the highest fraction of low-mass H I objects. For a comparison, HIPASS detects 41 objects (less than 1%) with H I masses below $10^8\ h_{70}^{-2}\ M_\odot$ in the whole southern sky $\delta < +2^0$.

We use the Schmidt (1968) method to derive the H IMF, taking into account the incompleteness of the survey for objects with small S_{int} values. The best-fit Schechter parameters of the H IMF which describes the H I mass distribution in the volume

probed by the WSRT CVn survey are $\alpha = -1.17$, $\log(M^*_{\rm HI}h^2_{70}/M_\odot) = 9.57$ and $\phi^* = 0.125h^3_{70}$ Mpc^{-3} dex^{-1}.

The slope α changes very little due to uncertainties in the adopted distances, measured fluxes or small number statistics of objects with high H I masses. It is accurate down to an H I mass of $\log(M^*_{\rm HI}h^2_{70}/M_\odot) = 6.40$ (lower limit of the lowest–mass bin). This is the lowest H I mass reached compared with any of the previous H IMF studies.

Acknowledgements

I would like to thank Saleem Zaroubi for kindly providing the grid of the peculiar velocities reconstructed from the surface brightness fluctuation survey of galactic distances (Tonry et al. 2001). This work greatly benefited from discussions with Rien van de Weygaert, Bernard Jones, Edwin Valentijn, Emilio Romano-Díaz and Julianne Dalcanton.

The Westerbork Synthesis Radio Telescope is operated by the ASTRON (Netherlands Foundation for Research in Astronomy) with support from the Netherlands Foundation for Scientific Research (NWO).